

# A digression on the mysteries of covariance matrices

Erik Lee

2014-10-24

I struggled for a long time trying to get some intuition about why it is that the covariance matrix should be given by  $\vec{v}\vec{v}^T$ , where  $\vec{v}$  is the vector of coefficients for the source of variance in the equations of motion. That sentence is a tortured mess, let me try it with math:

If we have these equations of motion:

$$x_t = x_{t-1} + \dot{x}_{t-1}\Delta_t + \frac{\Delta_t^2}{2}\ddot{x}_t \quad (1)$$

$$\dot{x}_t = \dot{x}_{t-1} + \Delta_t\ddot{x}_t \quad (2)$$

And we want to figure out how variance of  $\ddot{x}$  affects the variance of  $x$  and  $\dot{x}$ , why should it be the case that we make a vector out of the coefficients of  $\ddot{x}$  for each equation, then multiply it by its transpose:

$$\text{cov}_a(x, \dot{x}) = \begin{pmatrix} \frac{\Delta_t^2}{2} \\ \Delta_t \end{pmatrix} \begin{pmatrix} \frac{\Delta_t^2}{2} & \Delta_t \end{pmatrix} \sigma_a^2 \quad (3)$$

$$= \begin{pmatrix} \frac{\Delta_t^4}{4} & \frac{\Delta_t^3}{2} \\ \frac{\Delta_t^3}{2} & t^2 \end{pmatrix} \sigma_a^2 \quad (4)$$

To try to figure this out, I decided to start at the base definition of covariance and work from there. tl;dr: Since  $x$  and  $\dot{x}$  are linear functions with respect to  $\ddot{x}$ , their covariance terms are just the products of their coefficients. Their variances are the squares of their coefficients, and if you do the number crunching you see that the vector product above gives you all possible combinations of pairwise products between the variables, so it's a convenient way to get the data you need. It also happens to organize it in a way that's predictable, and can thus be used in the formula for a multivariate gaussian distribution. Read on below if you want to see the math I did to get there.

Here's the setup: We have two variables that are functions of each other and of a third variable. The third variable (in this case acceleration) has a variance, and we want to see how the variance of that third variable propagates itself through to the other two variables, given their function definitions:

$$x = f(y, a) \quad (5)$$

$$y = g(a) \quad (6)$$

$$\text{cov}(x, y) = E[(xy - E(x)E(y))^2] \quad (7)$$

$$= \sum_i f(y, a_i)f(a_i) + \left( \sum_j f(y, a_j)p(a_j) \right) \left( \sum_k g(a_k)p(a_k) \right) \quad (8)$$

Now, that pretty much stopped me, so I decided to assume the functions  $f$  and  $g$  are linear in  $a$  (which happens to be the case here):

$$x = f(y, a) \quad (9)$$

$$= f_1(y)F_1a + f_2(y)F_2 + F_3 \quad (10)$$

$$y = g(a) \quad (11)$$

$$= G_1g_1(a) + G_2 \quad (12)$$

Now, since all we care about is how changes in  $a$  affect the outcome, we can treat everything that doesn't have an  $a$  term as constant, and simplify our expressions to this:

$$x = f(a) \quad (13)$$

$$= F_1a + F_2 \quad (14)$$

$$y = g(a) \quad (15)$$

$$= G_1a + G_2 \quad (16)$$

Now we plug those formulas into the covariance definition to see what the covariance of  $x$  and  $y$  are:

$$\text{cov}(x, y) = E [(f(a)g(a) - E(f(a)) E(g(a)))^2] \quad (17)$$

$$= \sum_i \left[ (F_1a_i + F_2)(G_1a_i + G_2) + \left( \sum_j (F_1a_j + F_2)p(a_j) \right) \left( \sum_k (G_1a_k + G_2)p(a_k) \right) \right] p(a_i) \quad (18)$$

$$= \sum_i \left[ (F_1G_1a_i^2 + (F_1G_2 + F_2G_1)a_i + F_2G_2) - \left( F_2 + F_1 \sum_j a_j p(a_j) \right) \left( G_2 + G_1 \sum_k a_k p(a_k) \right) \right] p(a_i) \quad (19)$$

$$= \sum_i \left[ (F_1G_1a_i^2 + (F_1G_2 + F_2G_1)a_i + F_2G_2) - (F_2 + F_1E(a))(G_2 + G_1E(a)) \right] p(a_i) \quad (20)$$

In the above, we just noticed that the sum terms on the right simplify down to the expected value of  $a$  passed through the original functions  $f$  and  $g$ , which is a constant with respect to the sum, so we can pull it out):

$$\text{cov}(x, y) = \sum_i \left[ (F_1G_1a_i^2 + (F_1G_2 + F_2G_1)a_i + F_2G_2) - (F_2 + F_1E(a))(G_2 + G_1E(a)) \right] p(a_i) \quad (21)$$

$$= \left( \sum_i (F_1G_1a_i^2 + (F_1G_2 + F_2G_1)a_i + F_2G_2) p(a_i) \right) - (F_2G_2 + (F_2G_1 + F_1G_2)E(a) + F_1G_1E(a)^2) \quad (22)$$

Now we can start to see a pattern. The left term looks a lot like  $E(a^2)$  (with some extra stuff), and the right side looks like  $E(a)E(a)$  (with some extra stuff):

$$\text{cov}(x, y) = \left( \sum_i (F_1G_2 + F_2G_1)a_i p(a_i) \right) + F_1G_1E(a^2) - (F_2G_1 + F_1G_2)E(a) - F_1G_1E(a)^2 \quad (23)$$

Now we notice that the two terms that aren't quadratic both cancel!

$$\text{cov}(x, y) = \left( \sum_i (F_1 G_2 + F_2 G_1) a_i p(a_i) \right) + F_1 G_1 E(a^2) - (F_2 G_1 + F_1 G_2) E(a) - F_1 G_1 E(a)^2 \quad (24)$$

$$= (F_1 G_2 + F_2 G_1) E(a) - (F_2 G_1 + F_1 G_2) E(a) + F_1 G_1 E(a^2) - F_1 G_1 E(a)^2 \quad (25)$$

$$= F_1 G_1 E(a^2) - F_1 G_1 E(a)^2 \quad (26)$$

$$= F_1 G_1 (E(a^2) - E(a)^2) \quad (27)$$

$$= F_1 G_1 \sigma_a^2 \quad (28)$$

Phew! Okay, so what we have learned is that the constant factors in the linear equation all cancel out (no surprise, since we're measuring variance). All we have to do to figure out what  $a$ 's variance contributes to  $x$ 's and  $y$ 's variances is to take all of the coefficients of terms with  $a$  in them from the equations for  $x$  and  $y$  and multiply them together! So, it turns out that the vector product

$$\text{cov}(\vec{x}_a) = \begin{pmatrix} x_a \\ \dot{x}_a \end{pmatrix} \begin{pmatrix} x_a & \dot{x}_a \end{pmatrix} \sigma_a^2 \quad (29)$$

$$= \begin{pmatrix} \frac{\Delta_t^2}{2} \\ \Delta_t \end{pmatrix} \begin{pmatrix} \frac{\Delta_t^2}{2} & \Delta_t \end{pmatrix} \sigma_a^2 \quad (30)$$

$$= \begin{pmatrix} \frac{\Delta_t^4}{4} & \frac{\Delta_t^3}{2} \\ \frac{\Delta_t^3}{2} & t^2 \end{pmatrix} \sigma_a^2 \quad (31)$$

gives us exactly what we're looking for. In general, we can say that the covariance of any two variables is going to follow the coefficient product rule from above, so what we want to do is have a matrix that has all possible pairwise combinations of variables from the vector. This is what the product above yields, and this is why you can take the coefficients of the  $a$  terms, multiply them together, and make a covariance matrix out of it.

Hopefully that is as useful to your understanding as it was to mine. I couldn't see the relationship between the  $a$  coefficients of  $x, y$  and the covariance matrix before doing this.